

CORPUS-ASSISTED CREATIVE WRITING: INTRODUCING INTERMEDIATE ITALIAN LEARNERS TO A CORPUS AS A REFERENCE RESOURCE

Claire Kennedy and Tiziana Miceli

Griffith University

In much of the literature on the exploitation of corpora for language learning, the learners are viewed as researchers, who formulate and test their own hypotheses about language use. Having identified difficulties encountered in corpus investigations by our intermediate-level students of Italian in a previous study, we have designed a semester-long apprenticeship in corpus use which does not demand of them the high level of language proficiency, attention to detail in observation, and logical rigour that we consider necessary for rewarding work in the learner-as-researcher role. Instead, we introduce a corpus initially as an aid to the imagination in writing, and then to achieving accuracy through specific grammatical problem solving. We see this as the groundwork for subsequent development of the students' research skills with corpus data. This paper describes the approach we have adopted to the corpus apprenticeship and reports on an evaluation of its effectiveness through case studies of three students and their use of a corpus and bilingual dictionary as reference resources when writing. Drawing on insights from the case studies, we outline a working definition of corpus-consultation literacy for our learning context and identify some refinements to be made to our apprenticeship.

INTRODUCTION

Corpora and concordancers are now widely recognized as significant resources for L2 teaching and learning. Whether the learners have direct access to a concordancer or a teacher selects concordances for them to work on—approaches labelled as “hard” and “soft,” respectively, by Gabriellatos (2005)—there is broad agreement that working with corpus data can benefit learners by developing a process-oriented view of language learning and fostering their active involvement (O’Sullivan, 2007). Other key theoretical arguments for encouraging learners to examine concordances are that this can enhance their appreciation of the relationship between form and meaning, and therefore of the idea of grammar as essential to the construction of a text’s meaning (Gavioli, 1997), and stimulate attentiveness to the relationships between words and their context (Aston, 2001; Milton, 1999).

The grammatical consciousness-raising effects of work with concordances have been espoused particularly by Johns (1991), whose metaphor of the “learner-as-researcher” has proved very influential. Johns describes the learner armed with a corpus as a linguistic researcher who formulates and tests hypotheses on the basis of evidence from real language use. Thanks to the direct access to language data, the learner-researcher can engage in an individual process of discovery of grammar and meaning and take on greater responsibility for learning, while the teacher assumes the role of research coordinator rather than fount of knowledge. But Johns (1988, p. 21) also notes the potential for “serendipity learning,” or incidental discoveries in the course of corpus investigations. Bernardini builds on this in developing the “learner-as-traveller” approach to using corpora, where what matters is not the starting or end point of an investigation but the choices made, the strategies adopted, the experience gained and the findings along the way (2000, p. 142).

While many recent empirical studies have shown that corpus use can be highly productive and enjoyable for learners in various contexts, these generally positive results are sometimes “tempered in no small measure by the strongly worded negative reactions from the learners” (Chambers, 2007, p. 7). Some learners perceive corpus consultation as too time consuming (Chambers & O’Sullivan, 2004; Yoon & Hirvela, 2004) or laborious and tiring (Chambers, 2005; Cheng, Warren, & Xun-feng, 2003). Braun (2007) and Davies (2004) identify specific difficulties students face in managing a hypothesis-testing

process and in their observation and analysis of concordance data, while Granger and Meunier (2008, p. 248) note the challenge of motivating students to work with multi-word patterns if they perceive learning single words as already sufficiently demanding.

Several scholars therefore stress the importance of appropriate training in corpus consultation, through teacher-guided activities, before students embark on independent work (Chambers, 2005; Cheng et al., 2003; O'Sullivan & Chambers, 2006; Vannestal & Lindquist, 2007; Yoon & Hirvela, 2004). Bernardini (2004, p. 26) specifically recommends starting with “convergent” tasks—that is, tasks in which anyone working with the same data should arrive at the same outcome—so that the teacher can ensure certain basic techniques are applied and a correct path is followed through the processing of the data, before proceeding to “divergent” tasks, where different people working on the same data autonomously are expected to arrive at different outcomes (see also Leech, 1997). Frankenberg-Garcia (2005), in light of her analysis of students’ choices of reference resources for various types of queries, highlights the importance of presenting a corpus as an addition to the learners’ suite of resources rather than in isolation. With reference to this wealth of empirical studies, in a theoretical discussion of the assumptions underlying learners’ use of corpora and their implications for pedagogy, O’Sullivan (2007) calls for work to define corpus-consultation literacy and paths to achieving it.

We began to incorporate work with corpora into our Italian programme at a university in Australia several years ago, attracted by its potential usefulness in our context, where we seek to marry attention to grammatical form with a basically communicative teaching approach. Opting for a “hard” approach to corpus use by our students made sense in light of our emphasis on cultivating learning strategies to support their development as independent life-long learners. We consider appropriate training in corpus consultation to be especially important in our case: as FL learners in Australia, with few opportunities for immersion, our students are unlikely to have the high proficiency levels of the ESL, EAP and translation students referred to in many studies of corpus use; in fact, they typically reach what we consider a high-intermediate level after three years.

In this article we first illustrate our current approach to training our students in corpus consultation. We then present an evaluation of that approach through case studies of three students’ use of a corpus and dictionary as reference resources while writing. In discussing the insights gained from this empirical work, we seek to contribute to understanding of what constitutes corpus-consultation literacy and to support O’Sullivan’s theoretical case for giving greater prominence to corpus-consultation literacy in language learning. We also identify some refinements to be made to our training.

OUR CORPUS APPRENTICESHIP

Underlying our current approach to training in corpus use are three decisions, based on our experience with several cohorts of students and our response to the literature referred to above. First, we see mastering corpus consultation as a gradual, long-term process that needs to be treated as an integral part of the overall language-learning process. For this reason we decided to integrate training in corpus use fully into the curriculum—through a series of activities in class and for homework—and to look upon the learners’ difficulties and possibly negative responses as the kind of wavering to be expected in the early stages of a journey. Indeed, we call our semester-long introduction to corpus work an “apprenticeship,” in order to reflect this long-term perspective.

Second, we decided to insert the apprenticeship into a major creative writing project, this being the component of our Italian programme where we thought it most likely to engender a positive response. Central to our second-semester, second-year course is an autobiographical writing task, in which the students compose six chapters on set themes, published in installments throughout the semester on the course Website. They tend to become very enthusiastic about this task, as it brings the opportunity to write about topics that matter to them and a sense of achievement through sharing what they produce. We hoped their engagement with this project would help make them well disposed towards a corpus presented explicitly as a resource to support it. But we also saw the autobiography as a type of task within which we could make the usefulness of a corpus particularly apparent. It lends itself to addressing the following challenges discussed by Granger and Meunier (2008) and Milton (1999): raising learners’

consciousness about the importance of phraseology and encouraging them to venture beyond their existing repertoire of multi-word patterns. Furthermore, by virtue of its creative nature there is considerable scope for adapting the content to take advantage of the serendipitous discoveries of language patterns that a corpus facilitates.

Our third decision was to downplay the learner-as-researcher notion. Having identified difficulties encountered in corpus investigations by our students in a previous study (Kennedy & Miceli, 2001), we now postpone to after the apprenticeship the kind of corpus work that demands the high proficiency in understanding examples, and stringent observation and reasoning, that we consider inherent in a learner-as-researcher role and possibly responsible for some of the negative reactions reported in other studies. By this we refer to the kinds of investigations aimed at deriving a general and complete rule, such as answering the question “What is the difference between convince and persuade?” in the highly effective way that EAP students working with Johns (1991) were able to do. Such investigations are challenging for our intermediate-level students because they entail a thorough and detailed classification of the examples found in a search, and therefore require confidence in interpreting numerous examples and identifying small differences between them, in order to formulate any exceptions along with the general rule. They can also demand a high degree of rigour in observation and reasoning—research skills that our younger students are usually only in the process of developing in their studies in general. We therefore focus our apprenticeship instead on corpus use in relation to the specific requirements of a writing task to hand; that is, on investigations aimed at finding models for patterns to use at a given point in a given text.

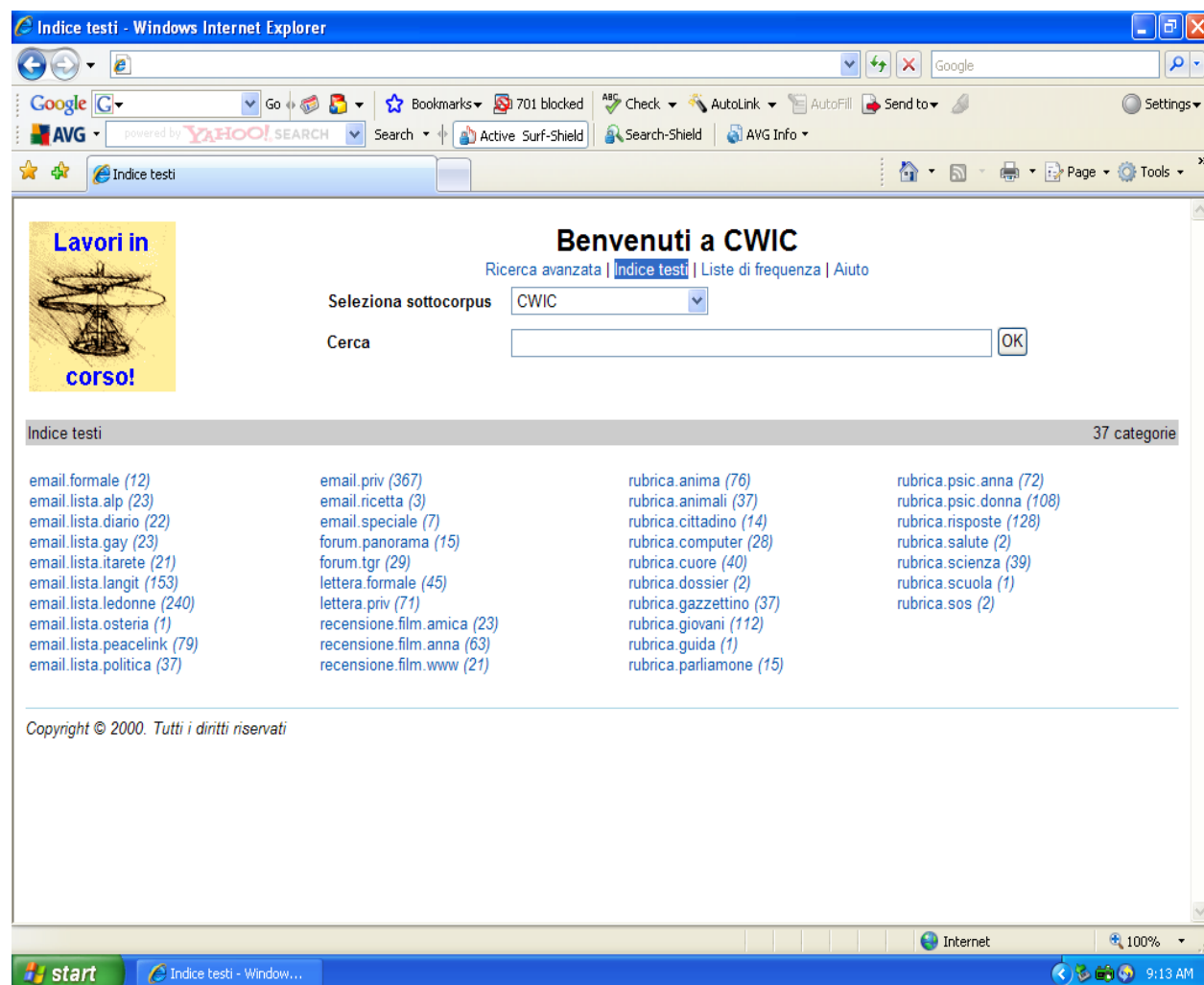


Figure 1a. Index of text-types present in CWIC, showing the number of texts included of each type.

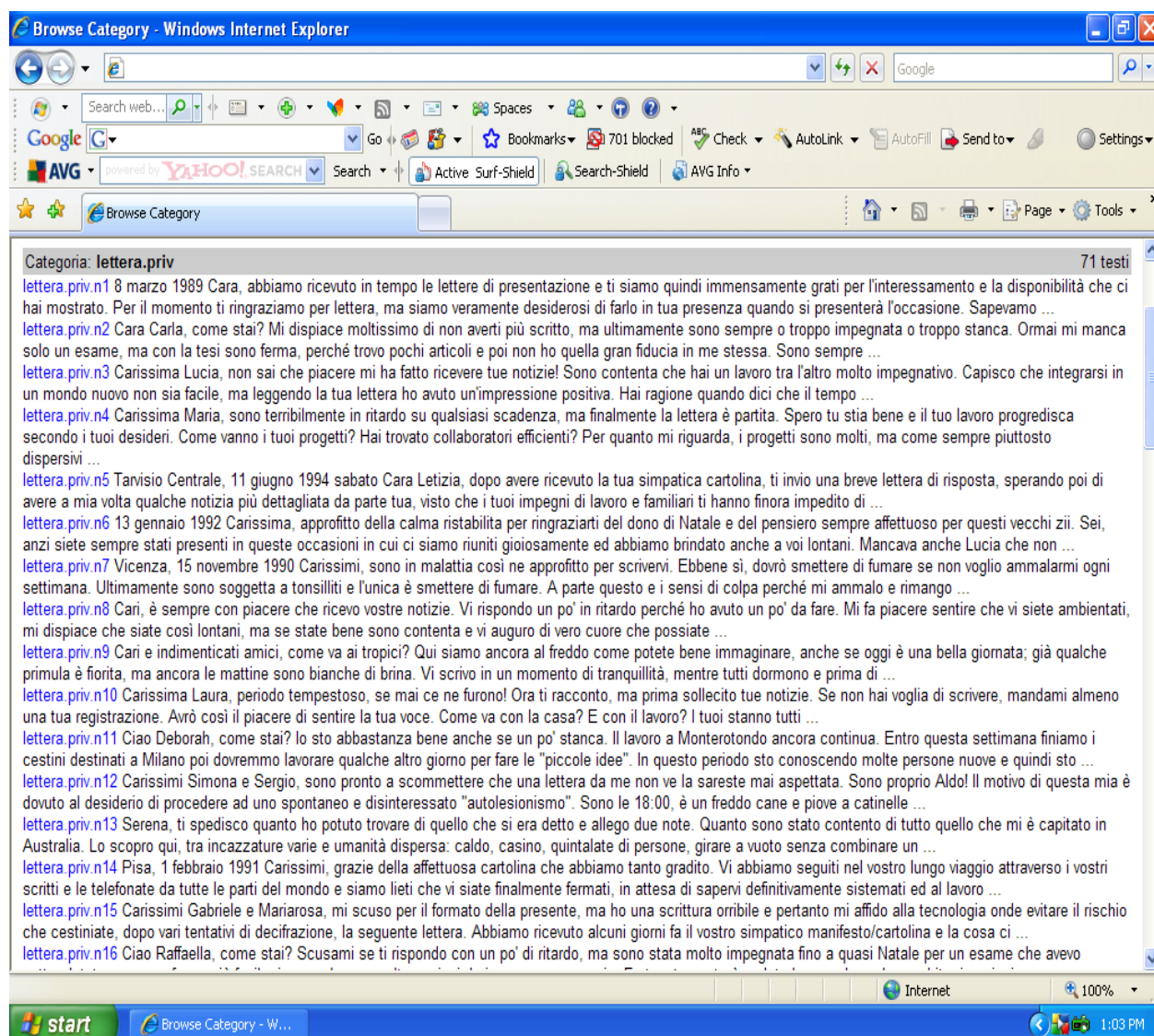


Figure 1b. List of texts of a single type (in this case personal letters), obtained by clicking on the name of the type (*lettera priv*) in the screen shown in Figure 1a.

The tools we use in the apprenticeship were custom-built for our students and learning context. We compiled the corpus, called CWIC (*Contemporary Written Italian Corpus*), a 500,000 word collection of letters, emails, and material from magazines, intended to provide models of personal writing on everyday topics, as well as texts by professional writers such as journalists and film critics. It can be interrogated via the Web or in standalone mode on a CD-ROM. The user interface was designed to look familiar to anyone accustomed to Internet searching.¹ In addition to concordancing, the search engine allows swapping between a concordance line and the whole text it came from (Figures 2a–b) and browsing whole texts by text type (Figures 1a–b), and produces frequency lists. We are fortunate in having access to a classroom whose configuration is particularly suited to the writing workshops and the corpus apprenticeship embedded in them, with computers handy and plentiful but not intrusive, as recommended by Chambers and Bax (2006, p. 471) for classroom CALL in general.

The thrust of the apprenticeship is to introduce the students to two key ways of using CWIC to enhance their writing: to enrich the content and language of their text, through what we call *pattern-hunting*; and to edit their text for lexico-grammatical accuracy, through *pattern-defining*. Both the pattern-hunting and

pattern-defining functions entail exploring the corpus in search of models for word patterns² to employ in one's own text (adapted as necessary), but their departure points differ.

Introducing the students to pattern-hunting amounts to encouraging them to use the corpus as an aid to the imagination and memory. All writers, and L2 learner writers in particular, can find themselves at a loss for ideas on what to write about and/or words to express it. In both cases the problem can be tackled by searching the corpus on words likely to be associated with the topic concerned, and then scanning the concordance lines for potentially useful patterns. For example, when the students were writing about their sense of personal space in one of their autobiography chapters, we suggested pattern-hunting in relation to questions such as "What sorts of things might I talk about in relation to my living space?" and "What adjectives or set phrases might I use to talk about my need for space?" Most started by searching on the familiar word *spazio*. This not only turned up ideas and expressions to borrow, such as "*reagire di fronte alle invasioni del proprio spazio*" (react to the invasion of one's space), "*ritagliare uno spazio per sé*" (carve out a space for oneself), and "*rubare spazio*" (take space)—see Figure 2a, lines 65, 57, 76—but also triggered further searches, on words encountered in the concordance lines, such as "*percorso*" (path) and "*compromesso*" (compromise). The other two principal techniques we introduce for pattern-hunting are browsing through whole texts, chosen on the basis of text type and title, and perusing frequency lists for common two-word, three-word, or four-word combinations.

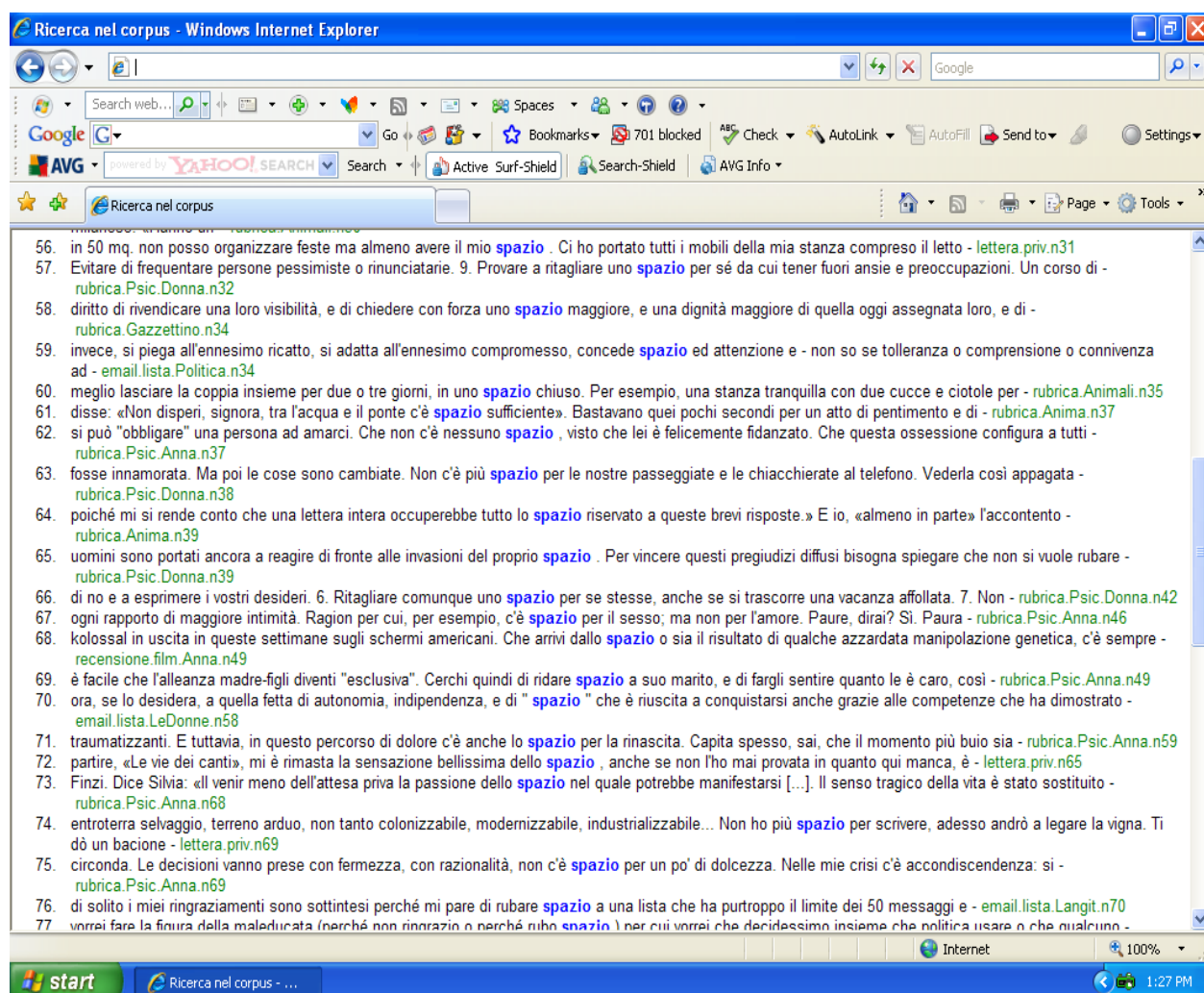


Figure 2a. Some results of a search on “*spazio*” used in a pattern-hunting operation when writing about personal space.

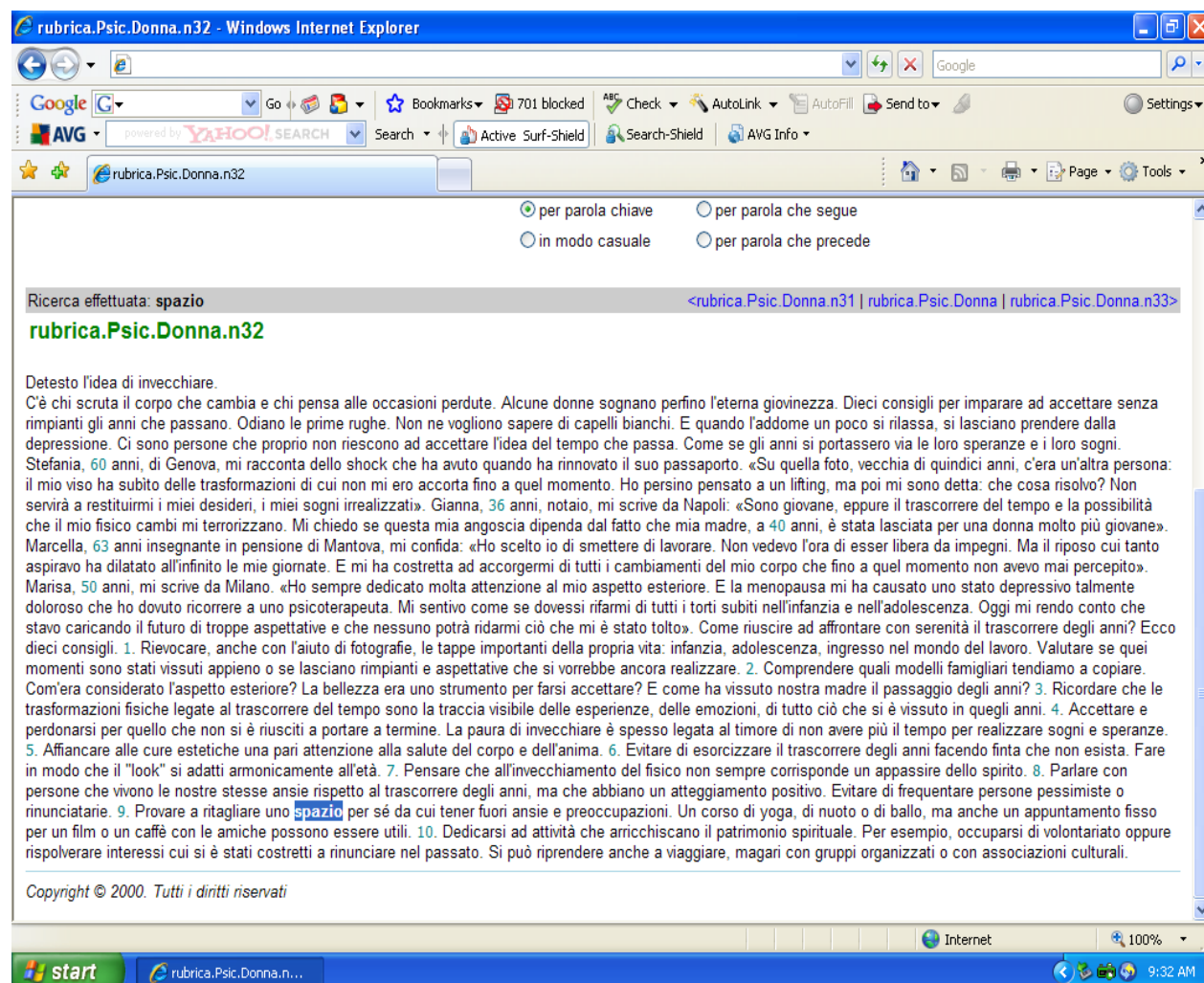


Figure 2b. Display of full text obtained by clicking on a concordance line (in this case from concordance line 57 in the screen shown in Figure 2a; the text is no. 32 of the letters to the “Psychology” column in the magazine *Donna*).

Pattern-hunting is somewhat similar to Milton’s (1999) wordlist-driven approach, intended to “[encourage] novice writers to experiment with grammatical and lexical patterns that they might otherwise be unaware of, or misuse, or avoid” (p. 243). We see both pattern-hunting and Milton’s approach as useful not only because they provide access to ways of saying things in the target language for which there may be no exactly corresponding pattern in the learners’ first language, but also because they can therefore highlight for the learners the significance of this non-correspondence phenomenon. Milton’s students work with a monolingual corpus enhanced by the provision of categorized lists of patterns that are deemed useful for the learners and whose usage can be explored in the corpus. The difference between the approaches is that Milton’s is focused on finding words to express what the user already has in mind to say (e.g., through a linking expression), while, as discussed above, pattern-hunting encompasses not only searching for *how* to say something but ideas on *what* to say about something. The different approaches reflect different orientations of the corpus use—to support creative writing in our case and academic writing in Milton’s.

The second function we introduce our students to, pattern-defining, differs from pattern-hunting in that it is concerned with finding models when we *do* have a specific target pattern in mind for use at a particular point in a text. Usually it is a matter of knowing some of the component words and seeking a model for the exact structure required; for example: which preposition is required after a verb in a certain context, or

the position of an adjective relative to a noun in a particular usage. Figure 3 provides an example of a pattern-defining operation carried out by one of our students.

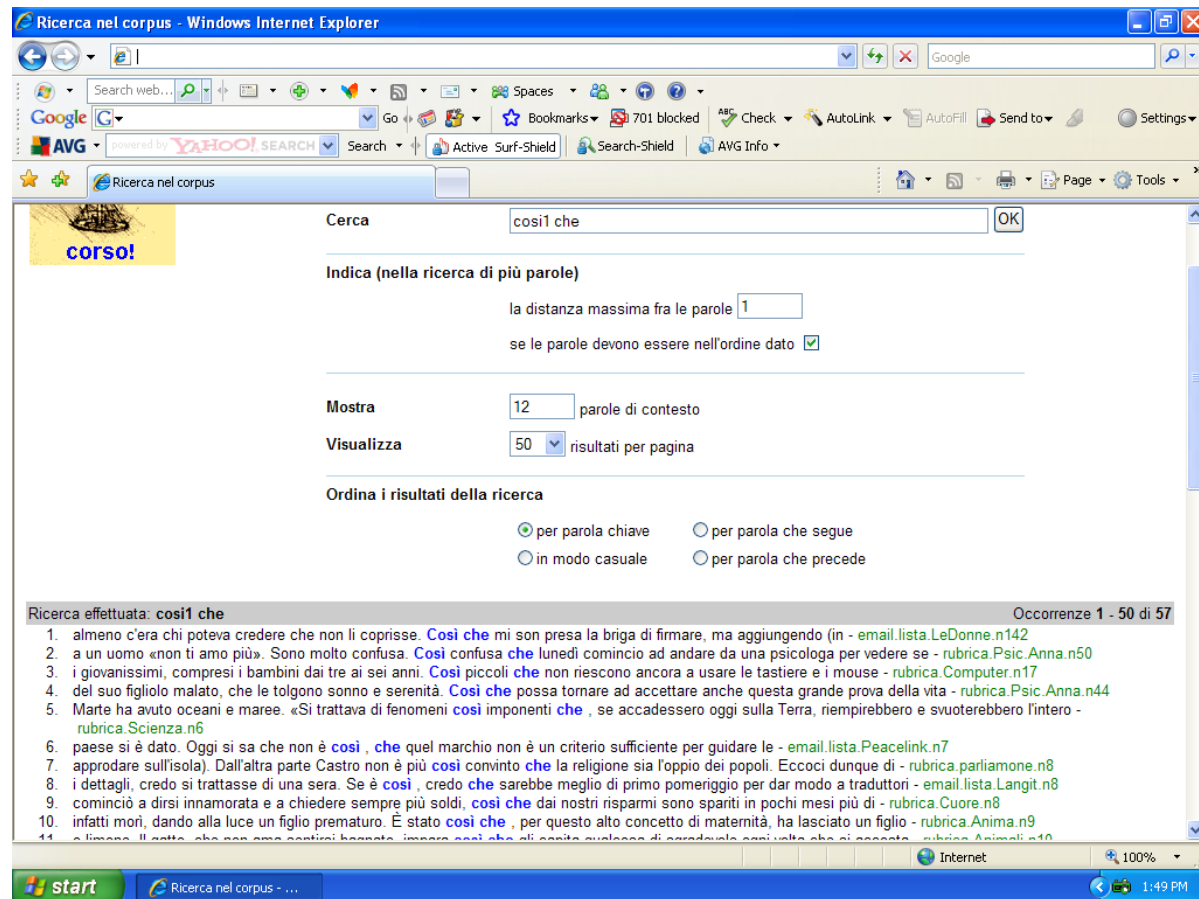


Figure 3. A search conducted by S3 in a pattern-defining operation aimed at establishing if the pattern “so” <adjective> “that” can be rendered in Italian with “così” <adjective> “che” and if the subjunctive mood is required after the “che.” The following parameters have been selected: at most 1 word distance between the two searched words, searched words only in the order given, 12 words of context displayed on each side, 50 concordances per page in random order.

We chose to concentrate on pattern-hunting and pattern-defining in the apprenticeship in the expectation that these functions can be productive and rewarding for learners even at an intermediate level of proficiency. We hope thus to acquaint our students with corpora as reference resources in a way that is motivating and stands them in good stead for subsequently developing into fully-fledged *language-learners-as-researchers*. The advantage of beginning with pattern-hunting is that, in the students’ first encounter with the corpus, the emphasis is on an individual exploration of the data rather than achieving correct answers. While there is a progression from teacher-guided work to less guided work in each phase—pattern-hunting and pattern-defining—there is not a progression from convergent to divergent tasks such as described by Bernardini (2004). Pattern-hunting work is divergent by definition: even when starting out with the same lookup, students will take different paths in exploring the data and come up with different outcomes. The teacher’s guidance in the pattern-hunting phase lies in recommending *types* of questions to pose and types of techniques to employ (searching on topic words, browsing whole texts and frequency lists). However, our second, pattern-defining, phase does rely on convergent tasks, as we take all students through some illustrative pattern-defining operations together.

EVALUATION OF THE APPRENTICESHIP AND IMPLICATIONS FOR DEVELOPING CORPUS-CONSULTATION LITERACY

Aims and method

At the end of the course in which we conducted the apprenticeship in its current form for the first time, we carried out an investigation into our students' use of CWIC and bilingual dictionaries, through case studies of three individual students. Our attention was focused on these two types of resources because they were the ones we had explicitly given training in, during that course and the prerequisite course in the previous semester, respectively. In investigating which functions the students used each resource for and how effectively, especially in light of Frankenberg-Garcia's (2005) finding that her students of translation favoured bilingual and mediated reference resources over monolingual and unmediated ones,³ we particularly sought to understand the ways in which the students had taken up CWIC—a monolingual and unmediated resource—as a result of our apprenticeship. We were especially interested in whether or not they had embraced both the functions of the corpus that we had concentrated on: pattern-hunting and pattern-defining.

Through the insights from the case studies we hoped to better understand what it should mean in our context to seek to develop students' corpus-consultation literacy, and thus identify ways of improving our approach to integrating the corpus into the students' suite of reference resources. The principal questions addressed were:

1. Which functions did the students use the corpus and dictionary for, and how effectively?
2. What factors, especially skills, knowledge and attitudes, affected their propensity and ability to use these reference-resource functions?
3. And therefore, how can we improve the apprenticeship to cultivate appropriate skills, knowledge, and attitudes?

The three students, whose participation in our study was voluntary, were taking a major in Italian within the undergraduate degree programme in Languages and Applied Linguistics. They were all hard-working and highly motivated, and had achieved grades of Distinction or High Distinction in their three previous Italian courses.⁴ None of the three had had any exposure to corpora prior to our course.

We do not claim the three students constituted a representative sample of their cohort but we consider them appropriate and interesting case-study subjects because of the different approaches to language learning and reactions to the corpus they exhibited in class. Furthermore, they came from different age groups and study experiences: two were women, aged 20 (S1) and 55 (S2) respectively, and the third a man aged 27 (S3). Two of them had had prior experiences of learning a language to a high proficiency level: S1 in English as a second language, learnt at high school in Australia; and S3 in Portuguese through formal instruction and immersion, for work purposes.

The primary method we used to obtain a detailed picture of the three participants' use of the resources was to set them a writing activity to work on independently for 45 minutes and interview them each for 45 minutes immediately afterwards. The purpose of the semi-structured interview was twofold: first, to extract a retrospective report on their use of reference resources during the writing activity, with a recording of their computer screen activity available to prompt the memory where necessary; and second, to discuss their habits, preferences and expectations regarding reference resources in general, and collect any feedback they offered on the apprenticeship. The interviews were audio-recorded and transcribed.

The writing activity we set was intended to give the participants ample scope for using the corpus in the ways practised in the course, as well as any other reference resources they habitually used. We invited them to reread the first and second chapters of their autobiography (written three months earlier), choose one of them and work on enriching it through additions and modifications. We stressed that this meant enhancing the text rather than correcting it grammatically, and that we wanted them to pay special attention to a suitably varied choice of vocabulary.⁵ We encouraged the participants to use any reference resources they chose. They brought their own printed bilingual dictionaries with them, and S2 also

brought a grammar book and a verb book. They used the computer we provided for access to CWIC and, in S3's case, an online bilingual dictionary. We also provided a printed monolingual dictionary, which none of them consulted; the word "dictionary" when used below therefore refers always to a bilingual dictionary.

In analysing each student's work during the case-study activity we classified the operations carried out by function⁶ and then type of resource, on the basis of the intentions declared in the retrospective report. The three functions we considered significant for the discussion below are shown in Table 1: pattern-hunting, pattern-defining, and finding an Italian equivalent for a given English pattern. It was important to include the function of finding an Italian equivalent—a classic function of the bilingual dictionary—because a few students had developed the habit of using the corpus to expand on the information found in the dictionary, even though this was not something that had been explicitly taught in class.⁷

Table 1. *Classification of Functions Used in the Case-Study Activity*

Function	Resource(s)
Pattern-hunting	corpus only
Pattern-defining	corpus only dictionary only dictionary and corpus (to expand on dictionary information)
Finding an Italian equivalent for a given English pattern	dictionary only dictionary then corpus (to expand on dictionary information)

We also collected data through an end-of-course questionnaire, on the students' perceptions of the autobiography task, and homework they completed in the second and sixth week of the apprenticeship with the corpus, respectively. The homework tasks concerned included suggestions intended to help the students practise using the corpus while writing their current autobiography chapter, in the ways they had been doing in class: pattern-hunting in the first week and pattern-defining in the fifth.

The discussion that follows is the fruit of our reflection on the qualitative data collected as described above. This reflection was aimed at describing each participant's use of CWIC and dictionaries, in the case-study activity and in general, and comparing their reference-resource use in light of their different perceptions of the autobiography task and the corpus apprenticeship. In this way we sought to identify factors that appeared to support or impede their taking up the various functions of the corpus, and that suggest practical steps to be taken to improve our training.

Results and discussion

There was marked variation among the three students in the frequency with which they used each resource in the case-study activity. The interviews confirmed that each had a different set of purposes to which he/she tended to put the resources and, in particular, had adopted a different set of functions of the corpus.

Participant S1

The young woman, S1, found the autobiography task particularly rewarding, both in a personal sense and for the development of her writing skills.⁸ Her motivation was enhanced by the knowledge that her work was to be published on the course Website and therefore "each instalment needed to be worth reading." In her writing she seemed prepared to take risks in sentence construction and word choice for the sake of expressiveness, where other students might restrict themselves to elements they were fairly confident of using correctly.

In the case-study activity, S1 was the only participant to use the corpus for both pattern-hunting and pattern-defining. Her aim in enriching the text was to convey more of how she felt about the friend who was the topic of the chosen chapter. She did this by adding elements of description and embellishing existing ones, usually inserting or replacing single words or groups of words rather than adding complete

sentences. She conducted 11 pattern-hunting operations, 9 of which led to borrowings; 1 successful pattern-defining operation using the corpus; 1 unsuccessful pattern-defining operation involving the dictionary and the corpus; and 3 dictionary searches for Italian equivalents, 2 of which were unsuccessful.

S1 was enterprising in her pattern-hunting, adopting various techniques for seeking out not just words but ideas on what to say. In addition to searching on a noun such as *esperienza* to find an adjective to describe an experience, or on a verb in order to borrow an adverb, she devised more ingenious strategies, including a search on *insieme* (together) for ideas on activities done together with her friend to relate, and a search on *mi manca* (I miss him) for help in adding what she missed in particular and how she felt. Another interesting strategy was searching on an adjective in order to find other adjectives, on the grounds that “when you look for one [there are] usually a few ... adjectives put together.” Also noteworthy was S1’s satisfaction with the spoils of her pattern-hunting: she evidently savoured some of the words and expressions found, such as *specchio di una profonda emotività* (reflection of a deep capacity for emotion), with reference to a person’s eyes. Some of her borrowings included the word(s) she had actually searched on, while others were of patterns found in the surrounding text: this was the case for the *specchio* expression, fruit of searching on *gli occhi* (eyes).

S1 seemed comfortable with CWIC’s monolingual and unmediated nature and not disconcerted by displays of large numbers of examples. Her approach to dealing with concordances would not have been sufficiently rigorous if she was in a learner-as-researcher role and needed to carefully classify examples in order to derive a general rule, but was usually perfectly adequate for the functions she used. She would scan the examples initially for any that appeared readily comprehensible and potentially useful, sometimes scrolling through several pages and considering only a few examples on each. If she found promising examples fairly quickly, she began to examine them more closely; otherwise she abandoned the operation.

While S1’s confidence in dealing with the concordances was usually warranted, and her attitude beneficial in that she was not deterred from experimenting with CWIC by lack of certainty of understanding them, it did mean she was likely to sacrifice concern for accuracy at times. Although she only once used a pattern from the corpus incorrectly in the case-study activity, it was clear from the retrospective report that she had sometimes been lucky, having not always paid sufficient attention to whether the words she borrowed were used in the corpus exactly as she intended to use them in her text. But it is worth noting that she made two similar errors in interpreting dictionary information, also due to too summary a consultation.

S1’s explanation of why she was comfortable with CWIC was that by practising on her own she had discovered through experience what she could do with it. In her independent work, she valued both intentional searches, aimed at finding patterns for including in specific sentences or exploring grammatical structures that had come up in class, and leisurely browsing with its potential for serendipitous discoveries. Asked when she had started to find CWIC useful, she explained:

...when I started using it at home and especially when we had to use it for our compiti [homework]... You find there’s more to it than is just on the first page you look up... Then there’s a lot of nice phrases that you can use, if not in that compito then in the next one or whatever you wish to write.... Sometimes you can look for a word or a phrase, or [sometimes] browse through all the articles; that’s when you’re not looking for anything in particular, you just want to learn new things... When I have time I just go and read things....

Clearly, S1’s positive reaction to the corpus was also facilitated by her strong engagement with the autobiography task through which she had been introduced to it, and by the fact that being online in the evenings—searching, browsing, chatting, or even trying out online translators—was a normal state for her.

Participant S2

In contrast to S1, S2 was only moderately engaged by the autobiography task, and CWIC evidently made little impression on her. She described the autobiography task as very demanding but a useful opportunity to develop her writing skills. Reflecting on her life story was not a new experience for her, as a middle-

aged person, and she did not see the task as drawing substantially on her imagination or affecting her personally. In the questionnaire she identified “dictionary work” as the area in which she had made most progress, through “striving to find the correct expressions to use.”

In the case-study activity, S2 used her (printed) dictionary for two significant and successful find-an-Italian-equivalent operations, and the corpus only once, for an unsuccessful pattern-defining operation. Her approach to enriching the text was not one that particularly lent itself to pattern-hunting: in inserting four new sentences at various places she focused on conveying exactly what she had in mind at the outset, in sentences as grammatically correct as possible. However, the main reason she did not do any pattern-hunting was, as emerged in the interview, that she had not adopted it as a corpus function at all. And her unsuccessful pattern-defining operation revealed a lack of expertise with the second function of the corpus. First, she failed to use the search parameters to implement her search in an efficient way, and therefore spent considerable time scrolling through irrelevant examples. Then, on (correctly) establishing that no useful examples were produced by that first search, she lacked the necessary experience and confidence either to persevere by trying substitutions (searching on a different but comparable adjective to the one she was interested in) or to choose to abandon the operation on the grounds that it might require too many attempts. She did abandon the operation, but this was because she lacked faith in her ability to use the corpus and feared wasting time, rather than because she judged CWIC might be inadequate for solving the particular problem in her limited timeframe.

However, it became clear in the interview that S2 was capable of great perseverance in dictionary use, where she was particularly expert and confident. She was prepared to spend considerable time and thought on a find-an-Italian-equivalent operation with her dictionary, trying various headwords and cross-checking between the English-Italian and Italian-English ends. Her dictionary expertise had developed over time, as dictionaries themselves had evolved, and she was completely familiar with the new features of her recently acquired dictionary.⁹ She was also aware of having mastered her grammar book gradually, despite initial frustration: “I used to hate the index in the grammar book.... It’s a matter of becoming familiar with something and using it and making it your own.”

“Making it her own” was what S2 had *not* done with the corpus: her understanding of its use(s) was not based on first-hand experience of any usefulness to her. In the interview, she reported that she had used it only when instructed to, in class activities and the relevant homework items. She identified two major deterrents. The first was CWIC’s monolingual nature: she considered her proficiency was not sufficient for her to “get the total meaning from the corpus” and saw great scope for getting things wrong if she failed to understand an array of examples completely. The second problem was her difficulty with the user interface: she felt she was slow in executing lookups, being confused by the search and display options available and hampered by being a “two finger typist.”

We identified a third impediment to corpus use for S2: that she had not appreciated the value of pattern-hunting as a function that she *could* use to good effect even at her level of proficiency. Yet she had pattern-hunted with evident satisfaction in class at the start of the apprenticeship, as she recalled when prompted: “Ah yes that was... pretty good: putting in the word... and seeing what came up. Particularly *amore*. And a couple of times I found expressions [and] thought ‘Oh I’ll use that in my autobiography.’” She seemed not to have understood that we were proposing this as a general strategy for the students to adopt, not just as a playful introduction to the corpus content.

In fact, ignoring pattern-hunting contributed to a more general misconception of S2: associating the corpus with the same functionality as the dictionary and grammar book. The ideas she expressed on the uses of the corpus, in contrast to those on the dictionary, were quite vague; for example: “I use the dictionary so much ... to look for, you know, those phrases and expressions ... That’s what I could come to CWIC looking for...” But given her sense of mastery of the dictionary and grammar book, it would clearly be pointless for her to seek answers in CWIC to problems she expected to solve more quickly and easily with them.

We drew lessons from S2’s case for our future practice. The first was that more ongoing assistance in practising pattern-hunting and pattern-defining, through homework instructions, would have been useful for her. The second was that students with lower Internet literacy may need extra training in using the

search-engine interface. While most of the other students drew on a well-developed intuition for operating online resources, for S2 the CWIC interface was something she had to learn to use. At the interview we realized that she was not familiar with some important features; for example, she had difficulty in restricting the numbers of examples displayed (by searching on combinations of words), which undoubtedly exacerbated her sense of being overwhelmed by the problem of understanding the examples.

Participant S3

The third case-study participant was like S1 in his propensity to seek out new words and expressions, while he shared with S2 a great concern for accuracy. For S3 the autobiography task was a good opportunity to develop his writing skills because, as he said “I always try to avoid simplistic writing and test ideas or ways I think things might be said, and the autobiography has been a good forum for that.” He also found it an interesting experience personally, in that it caused him to reflect on his life, but did not see it as an exercise in creative writing.

S3’s approach to the case-study activity reflected attitudes he conveyed later in the interview: first, his aim of expressing himself in written Italian with the same sophistication as in English; and second, his view of grammar as the essence of this, and the key to success in his previous language-learning experiences. In enriching the text he added a six-sentence paragraph, paying considerable attention to vocabulary choices and complexity of sentence structure, and also replaced a couple of adjectives in the existing paragraphs. In the process, he conducted eleven find-an-Italian-equivalent operations—five with the dictionary alone and six using the dictionary and CWIC—and was satisfied with the outcomes of all but two. He also used the corpus alone for two successful pattern-defining operations (see [Figure 3](#)), but not for pattern-hunting.

In the interview S3 stated that the bilingual dictionary “naturally remain[ed] the first point of reference” when searching for Italian equivalents to English words, but that CWIC had become his first choice for answering specific questions about lexico-grammatical patterns (i.e., pattern-defining). A characteristic of his reference-resource behaviour was the use of CWIC in find-an-Italian-equivalent operations, in order to expand on the dictionary information or seek corroboration of his interpretation of it. This technique had not been practised in class during the apprenticeship but he had developed it for himself very effectively. He listed several valid reasons for complementing a dictionary search with corpus searches:

If I don’t completely understand it from the dictionary but can see it in use in CWIC then it makes it a lot clearer.... [Or] I check to see how common it is¹⁰ or if I’ve interpreted it right...or if you need certain prepositions or conjunctions after it...to get a bit more of the context.

Although he was curious about words, and found the corpus appealing for that reason, S3 had not taken up pattern-hunting. In part it seemed that, like S2, he had not perceived it as a general strategy we were proposing, and he was not very inclined to use the corpus as an aid to his imagination anyway, because he focused on trying to convey in Italian exactly what would have said in English.

What was most striking about S3’s work with the corpus was the set of attitudes he expressed and manifested that evidently helped make him confident and his work effective. First, he saw mastering CWIC as a long-term, individual process for which he was responsible: “I wouldn’t always draw the right conclusion from what I’d found... I guess it’s something you gain through experience, as you go along....” Second, he appreciated that working with CWIC meant accepting the uncertainty of finding answers:

Sometimes you just can’t get your head around it. ... Sometimes things would just kind of leap out at you and other times you’d be muddling around with it for fifteen minutes and feeling like you’re getting nowhere... Grammar is so complicated, to find the right answers you need to know the right questions to ask of it. And... if you don’t understand it completely then it’s hard to know what the right questions are....

Third, he appeared to treat each unsuccessful operation as a specific case, rather than as an indicator of any general inadequacy on his part or that of the corpus. In the case-study activity he was not demoralized

or flustered when his searches failed to generate useful examples, or produced examples he was unable to interpret confidently.

Insights from the case studies regarding corpus-consultation literacy

The three participants used the resources in such different ways that we perceived them to have individual “reference-resource-using styles.” Each had become a sophisticated and confident user of her/his preferred resource(s) for selected functions. Notably, S1 and S3 had developed their own range of strategies for using CWIC in pattern-hunting and finding Italian equivalents, respectively. All three made the point that the essence of mastering CWIC was “making it your own,” and S1 and S3 both felt rewarded by having done so through experimenting with it independently.

Although we cannot pretend to generalize from these three cases to categories of reference-resource users, we find it highly suggestive that the two students who put a premium on correctness in their written Italian, and at conveying exactly what they set out to say in their autobiography (S2 and S3), engaged frequently in finding Italian equivalents and pattern-defining but not pattern-hunting, while the student who was most personally engaged with the autobiography task, and most inclined to see it as creative writing (S1), gave priority to discovering new patterns over solving predefined problems. S3’s attitude to grammar as what made language learning effective seemed to have predisposed him to appreciating the corpus as a tool for getting at the detail of the ways words combine to convey meaning, and therefore as a welcome addition to his resource kit for achieving sophisticated use of the language. For S2, by contrast, the exposure to the corpus in our apprenticeship had not sufficed to erode her equation of reference resources with the traditional functions of the dictionary and grammar book, so that the corpus remained an unnecessary extra.

Given the individual nature of these students’ approaches to reference resources, we must accept that we may not be able to convince all students to use CWIC in all the ways we would like. The case studies have nevertheless shown us ways in which we may be able to improve the apprenticeship so as to ensure the students have a better grounding on which to base their individual responses to the corpus. The principal practical finding is that we need to convey the value of pattern-hunting better, ensuring the students see it as a function in its own right, albeit not the type of function they may have previously associated with reference resources.

More generally, in analysing the three students’ work, we have come to see the process of “making the corpus your own” as a matter of developing not only skills in conceptualizing and executing searches and interpreting examples, but also appreciation of a set of principles that underpin effective use of the corpus and reference resources in general. In future, through class discussion around corpus-based activities we will seek to make these principles explicit, in order to help the students become more rigorous in their reference resource use, whatever their individual style.

The first principle is that we should think in terms of having a set of distinct reference-resource *functions* at our disposal, rather than simply a suite of reference resources, for what matters is not mastery of a resource but of the function(s) it can be used for. In future, we will stress that by taking up the corpus the students are extending their repertoire of reference-resource functions, to include one that is peculiar to the corpus (pattern-hunting), one that is also a function of the dictionary and grammar book but that can be carried out more efficiently with the corpus in many cases (pattern-defining), and one that sees it used in conjunction with the dictionary (finding-an-Italian-equivalent). We will seek to prevent their forming vague notions about the corpus as complementing the traditional resources, by labelling each corpus function as it is introduced, and encouraging the students to refer to all reference-resource functions by name so as to distinguish them from each other. Furthermore, we will dedicate more discussion and practice time in class to the importance of, and factors in, choosing the appropriate resource-and-function combination for the problem at hand.

Linked to this is the second principle: the importance of understanding the way the functions are linked to the characteristics of the resource. This will be especially important in supporting our greater emphasis on pattern-hunting: we will stress that the value of pattern-hunting derives precisely from the unmediated and monolingual nature of the corpus. We will seek to raise students’ awareness of both its practical value—

access to patterns we would not think to look for, or would not necessarily find, in a find-an-Italian-equivalent operation—and the significance of the underlying concept of non-correspondence between ways of expressing oneself in English and Italian.

The third principle is also a corollary of the first: that it is necessary to examine the limitations of reference resources, and one's own limitations in using them, with respect to specific functions, and specific problems they are used to address, rather than the resource itself.¹¹ All three case-study participants showed awareness that CWIC was too small to necessarily provide useful examples for any given problem, but their statements on this suggested there was a risk they might see *all* conclusions drawn from corpus operations as therefore tentative or incomplete, whereas we can have a high degree of confidence in answers obtained from CWIC to many types of questions. As far as users' limitations are concerned, failure to understand all the examples displayed by a corpus search might sabotage an investigation aimed at deriving a rule and the details of its exceptions, but in pattern-hunting it is perfectly legitimate to work only with examples that are easily understood, as the aim is not to find *all* relevant patterns. Where pattern-defining and finding an Italian equivalent are concerned, again it may not be necessary to understand many examples to obtain an answer: a small number can suffice as long as they provide close enough models to the case being investigated.

The fourth principle is that working with corpora requires greater preparedness to proceed by trial and error than work with other reference resources, and acceptance of the uncertainty of finding a satisfactory answer. Admittedly, acceptance of uncertainty can be difficult to acquire in many of life's circumstances, but at least we can aim to raise the students' consciousness, through discussion and practical examples, regarding the validity of giving up on an operation that seems to be getting nowhere. Abandoning a reference-resource operation has to be understood as a strategic decision, consistent with a trial-and-error approach to a problem, rather than as a failure.

The fifth principle is that the development of skills in corpus use is a long-term process, as it is for other reference resources. We teachers, as much as the students, need to acknowledge this, contain our expectations of the outcomes of a one-semester apprenticeship, and continue to actively support the development of these skills in subsequent courses. After all, our students' relative confidence in using bilingual dictionaries for finding Italian equivalents is the fruit of long experience: most have been using monolingual dictionaries for help with their native language since primary school and bilingual dictionaries in foreign language learning since high school.

We think it may accelerate the integration of CWIC into our students' learning environment if we try to enhance its legitimacy by drawing their attention to other Italian corpora and to the extensive use of English corpora around the world in ESL learning contexts. For it seems that the fact that CWIC was the only corpus they had ever encountered, combined with the knowledge that it was of our own creation, gave some students the impression it was a one-of-a-kind, experimental resource. Meanwhile, the legitimacy of dictionaries and grammar books in the students' eyes may be partly due to the existence of numerous brands, each providing the same functionality.

In addition to the various planned enhancements to our apprenticeship noted above, we intend to introduce an individual tutorial session for each student, towards the end of the semester. This will be similar to the interviews conducted with the case-study participants, which they found particularly valuable. At one level there are the practical benefits to students of clearing up misconceptions and being reminded of certain features and functions. But there are also the benefits that come from taking time to reflect on and discuss their use of the resources and attitudes to them that a one-to-one session allows.

CONCLUSION

These three case studies have provided a starting point for defining what is entailed in developing corpus-consultation literacy in our learning context. We now see it as a matter of promoting not only skills in using corpora for certain functions—at intermediate level: pattern-hunting, pattern-defining, and expanding on dictionary information in seeking an Italian equivalent—but also appreciation of certain principles underpinning effective reference-resource consultation and their implications.

The main thrust of the change we are making to our apprenticeship for the future is to work on raising the students' consciousness in relation to the five principles set out in the previous section, through discussion in class and in individual tutorial sessions. We will also be taking practical steps, again informed by the case studies: giving greater prominence to pattern-hunting, paying particular attention to technical training for the students who are not search-engine savvy, labelling reference-resource functions to help students recognize them as distinct, and further integrating work with the corpus into compulsory homework items in order to encourage the students to experiment for themselves. We will draw on the interviews with the case-study participants for specific insights to pass on to future cohorts, such as S1's techniques for pattern-hunting and S2's recommendation of a particular dictionary.

In light of our view of reference resources as providing an aid to the learners' imagination as well as to their achievement of accuracy, we have come to see our students in this apprenticeship as perhaps better represented by the metaphor of "learner as chef" than "learner as researcher" or "learner as traveller." An apprentice chef expects to develop expertise both from recipes and from the observation of master chefs' manipulations of ingredients and their outcomes, and seeks—from both sources—not only rules but stimuli to creativity. This is the overarching theme under which we will conduct our revised apprenticeship in future, and it will constitute a further key topic for discussion with the students.¹²

NOTES

1. The search engine was implemented by John Jeffery of Software Projects P/L, Brisbane, to whom we are extremely grateful.
2. The word "pattern" is used here to mean any kind of grouping of one or more words, including set phrases, idiomatic expressions and words in a relationship of collocation or colligation.
3. Frankenberg-Garcia (2005) used the term "unmediated" to refer to resources that require the user to interpret data, as opposed to "mediated" resources that provide the user with explanations or answers.
4. The possible grades for each course are: High Distinction, Distinction, Credit, Pass, Fail.
5. Our justification in proposing this as a worthwhile exercise was that the texts had been written, and corrected in response to the teacher's feedback, three months previously, in the first weeks of semester. Since then the students had not only been introduced to a new resource in the corpus but had greatly extended their command of Italian through writing the autobiography, reading short stories in the other strand of the same course, and the various components of the companion course in spoken Italian.
6. Our use of the term "function" differs from our use of "lookup" or "search" in that carrying out a single function may involve more than one lookup or search, possibly using more than one resource.
7. We excluded from consideration in the discussion the few quick lookups of a bilingual dictionary or the corpus to check spelling or gender, and verb books or dictionaries to find specific verb conjugations.
8. S1 said in the questionnaire: "I loved this assessment.... a chance to write something I wanted to.... I don't [usually] have a lot of time to reflect on my past.... I realized how much I miss my friends..." and later: "I really liked it because we get to use so many new words, we improve...our writing skills overall."
9. For example, this dictionary includes an indexed, bilingual supplement containing models of letters, emails, faxes, and phone call dialogues for various personal and business purposes from which S2 had borrowed—"copying bits and pieces"—in carrying out two authentic tasks she had set herself during semester: a letter to Italian acquaintances and an email to an Italian hotel to make an enquiry.
10. Evidently, for S3, a sufficient condition for a pattern to be considered as in common use, and therefore appropriate for his autobiography, was that it occurs a few times in CWIC. We find that reasonable as a rule of thumb, although checking the text type(s) it occurs in can be advisable too. We make no claim that

occurring in CWIC is a *necessary* condition for a pattern to be considered as in common use, but S3 did not seem to either.

11. Here we are leaving aside limitations in relation to access, which apply to the resource itself. These may be a matter of the user's individual perception. For example, for S2, a limitation of the corpus was that it is not portable like her printed dictionary. But the opposite view was held by several of our young students, who were more likely to be online, especially if engaged in writing something, than to have a printed dictionary on hand.

12. Our initial ideas on this theme were presented at the Fourth Biennial Conference of the Australasian Centre for Italian Studies in July 2007 (see www.acis.org.au/fourth_bienConf.html) and are being further developed in a paper currently in progress.

ACKNOWLEDGMENTS

We are most grateful to the three reviewers for valuable suggestions that led to significant improvements in this paper.

ABOUT THE AUTHORS

Claire Kennedy is Cassamarca Senior Lecturer in Italian Studies at Griffith University. Her current research interests are computer-assisted language learning and motivation in language learning.

E-mail: c.kennedy@griffith.edu.au

Tiziana Miceli is a Lecturer in Italian Studies in the School of Languages and Linguistics at Griffith University. Her research interests include computer-assisted language learning, including electronic corpora, blogs and wikis; language learning strategies; and language teaching methodology.

E-mail: t.miceli@griffith.edu.au

REFERENCES

- Aston, G. (2001). Learning with corpora: An overview. In G. Aston (Ed.), *Learning with corpora* (pp. 7-45). Bologna: CLUEB.
- Bernardini, S. (2000). *Competence, capacity, corpora. A study in corpus-aided language learning*. Bologna: CLUEB.
- Bernardini, S. (2004). Corpora in the classroom: An overview and some reflections on future developments. In J. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 15-36). Amsterdam: John Benjamins.
- Braun, S. (2007). Integrating corpus work into secondary education: From data-driven learning to needs-driven corpora. *ReCALL*, 19(3), 307-328.
- Chambers, A. (2005). Integrating corpus consultation in language studies. *Language Learning and Technology*, 9(2), 111-125. Retrieved from <http://llt.msu.edu/vol9num2/chambers/default.html>
- Chambers, A. (2007). Popularising corpus consultation by language learners and teachers. In E. Hidalgo, L. Quereda, & J. Santana (Eds.), *Corpora in the foreign language classroom. Selected papers from TaLC 2004* (pp. 3-16). Amsterdam: Rodopi.
- Chambers, A., & Bax, S. (2006). Making CALL work: Towards normalisation. *System*, 34(4), 465-479.

- Chambers, A., & O'Sullivan, Í. (2004). Corpus consultation and advanced learners' writing skills in French. *ReCALL*, 16(1), 158-172.
- Cheng, W., Warren, M., & Xun-feng, X. (2003). The language learner as language researcher: Putting corpus linguistics on the timetable. *System*, 31(2), 173-186.
- Davies, M. (2004). Student use of large, annotated corpora to analyze syntactic variation. In G. Aston, S. Bernardini, & D. Stewart (Eds.), *Corpora and language learners* (pp. 245-270). Amsterdam: John Benjamins.
- Frankenberg-Garcia, A. (2005). A peek into what today's language learners as researchers actually do. *International Journal of Lexicography*, 18(3), 335-355.
- Gabrielatos, C. (2005). Corpora and language teaching: Just a fling, or wedding bells? *TESL-EJ* 8(4), A1, 1-37. Retrieved from <http://www.tesl-ej.org/wordpress/>
- Gavioli, L. (1997). Exploring texts through the concordancer: Guiding the learner. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 83-99). London: Longman.
- Granger, S., & Meunier, F. (2008). Phraseology in language learning and teaching: Where to from here? In S. Granger & F. Meunier (Eds.), *Phraseology in foreign language learning and teaching* (pp. 247-252). Amsterdam: John Benjamins.
- Johns, T. (1988). Whence and whither classroom concordancing. In T. Bongaerts, P. De Haan, S. Lobbe, & H. Wekker (Eds.), *Computer applications in language learning* (pp. 9-27). Dordrecht: Foris.
- Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning materials. *English Language Research Journal*, 4, 1-16.
- Kennedy, C., & Miceli, T. (2001). An evaluation of intermediate students' approaches to corpus investigation. *Language Learning and Technology*, 5(3), 77-90. Retrieved from <http://llt.msu.edu/vol5num3/kennedy/default.html>
- Leech, G. (1997). Teaching and language corpora: A convergence. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and Language Corpora* (pp. 1-23). London: Longman.
- Milton, J. (1999). Lexical thickets and electronic gateways: Making text accessible by novice writers. In C. Candlin & K. Hyland (Eds.), *Writing: Texts, processes and practices* (pp. 221-243). London: Longman.
- O'Sullivan, Í. (2007). Enhancing a process-oriented approach to literacy and language learning: The role of corpus consultation literacy. *ReCALL*, 19(3), 269-286.
- O'Sullivan, Í., & Chambers, A. (2006). Learners' writing skills in French: Corpus consultation and learner evaluation. *Journal of Second Language Writing*, 15(1), 49-68.
- Vannestål, M. E., & Lindquist, H. (2007). Learning English grammar with a corpus: Experimenting with concordancing in a university grammar course. *ReCALL*, 19(3), 329-350.
- Yoon, H., & Hirvela, A. (2004). ESL student attitudes towards corpus use in L2 writing. *Journal of Second Language Writing*, 13, 257-283.